Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 584712, 6 pages
http://dx.doi.org/10.1155/2015/584712

Research Article

The Sparsity of Underdetermined Linear System via l_p Minimization for $0 < p < 1$

Haiyang Li,[1] Jigen Peng,[2] and Shigang Yue[3]

[1] *School of Science, Xi'an Polytechnic University, Xi'an 710048, China*
[2] *Department of Mathematics, Xian Jiaotong University, Xi'an 710049, China*
[3] *School of Computer Science, University of Lincoln, Lincoln LN6 7TS, UK*

Correspondence should be addressed to Haiyang Li; fplihaiyang@126.com

Received 7 November 2014; Accepted 14 April 2015

Academic Editor: Kacem Chehdi

The sparsity problems have attracted a great deal of attention in recent years, which aim to find the sparsest solution of a representation or an equation. In the paper, we mainly study the sparsity of underdetermined linear system via l_p minimization for $0 < p < 1$. We show, for a given underdetermined linear system of equations $A_{m\times n}X = b$, that although it is not certain that the problem (P_p) (i.e., $\min_X \|X\|_p^p$ subject to $AX = b$, where $0 < p < 1$) generates sparser solutions as the value of p decreases and especially the problem (P_p) generates sparser solutions than the problem (P_1) (i.e., $\min_X \|X\|_1$ subject to $AX = b$), there exists a sparse constant $\gamma(A, b) > 0$ such that the following conclusions hold when $p < \gamma(A, b)$: (1) the problem (P_p) generates sparser solution as the value of p decreases; (2) the sparsest optimal solution to the problem (P_p) is unique under the sense of absolute value permutation; (3) let X_1 and X_2 be the sparsest optimal solution to the problems (P_{p_1}) and (P_{p_2}) $(p_1 < p_2)$, respectively, and let X_1 not be the absolute value permutation of X_2. Then there exist $t_1, t_2 \in [p_1, p_2]$ such that X_1 is the sparsest optimal solution to the problem (P_t) $(\forall t \in [p_1, t_1])$ and X_2 is the sparsest optimal solution to the problem (P_t) $(\forall t \in (t_2, p_2])$.

1. Introduction

Recently, considerable attention has been paid to the following sparsity problem. Given a full-rank matrix A of size $m \times n$ with $m \ll n$, m-vector b, and knowing that $b = AX^*$, where $X^* \in \mathbf{R}^n$ is an unknown sparse vector, we expect to recover X^*. Although the system of equations is underdetermined and hence it is not a properly posed problem in linear algebra, sparsity of X^* is a very useful priority that sometimes allows unique solution. Accordingly, one naturally proposes to use the following optimization model (P_0) to obtain the sparsest solutions:

$$(P_0) \quad \min_X \quad \|X\|_0 \qquad \text{s.t.} \quad AX = b, \tag{1}$$

where $\|X\|_0$ denotes the number of nonzero components of X (we call $\|\cdot\|_0$ l_0 norm). This is one of critical problems in compressed sensing research. This problem is motivated by data compression, error correcting codes, n-term approximation, and so forth (see, e.g., [1]). It is known that the problem (P_0) needs nonpolynomial time to solve (cf. [2]). It is crucial to recognize that one natural approach to tackle (P_0) is to solve the following convex minimization problem:

$$(P_1) \quad \min_X \quad \|X\|_1 \qquad \text{s.t.} \quad AX = b, \tag{2}$$

where $\|X\|_1 = \sum_{i=1}^n |x_i|$ is the standard l_1 norm. The study of this problem (P_1) was pioneered by Donoho, Candès, and their collaborators and many researchers have made a lot of contributions related to the existence, uniqueness, and other properties of the sparse solution as well as computational algorithms and their convergence analysis to tackle the problem (P_0) (see survey papers in [3–5]). However, the solutions to the problem (P_1) are often not as sparse as those to the problem (P_0). It is definitely imperative and required for many applications to find solutions which are more sparse than that to the problem (P_1). A natural try for this purpose is to apply the problem (P_p) $(0 < p < 1)$, that is, to solve the following model:

$$\begin{aligned} (P_p) \quad \min_X \quad & \|X\|_p^p \\ \text{s.t.} \quad & AX = b, \end{aligned} \tag{3}$$

where $\|X\|_p^p = \sum_{i=1}^n |x_i|^p$ (we call $\|\cdot\|_p$ l_p-norm, though it is no longer norms for $p < 1$ as the triangle inequality is no longer satisfied). Obviously, the problem (P_p) is no longer a convex optimization problem. This minimization is motivated by the following fact:

$$\lim_{p \to 0_+} \|X\|_p^p = \|X\|_0. \tag{4}$$

This model was initiated by [6] and many researchers have worked on this direction [1, 2, 7–16]. They demonstrate that (1) for a Gaussian random matrix A, the restricted p-isometry property of order s holds if s is almost proportional to m when $p \to 0_+$ (cf. [8]); (2) when $\delta_{2s} < 1$ (or $\delta_{2s+1} < 1$, $\delta_{2s+2} < 1$), the optimal solution to the problem (P_p) is the same as the optimal solution to the problem (P_0) when $p > 0$ small enough, where $\delta_{2s} < 1$ is the restricted isometry constants of matrix A (similar for $\delta_{2s+1} < 1$, $\delta_{2s+2} < 1$) (cf. [7, 10, 13]); and (3) the l_p minimization can be applied to a wider class of random matrices A (cf. [11]). In addition, in [7, 15], the authors show that the problem (P_p) generates sparser solution than the problem (P_1) and the problem (P_p) generates sparser solution as the value of p decreases by taking phase diagram studies with a set of experiments. Nevertheless, are the conclusions showed by taking phase diagram studies true in theory? In the paper, we will answer this question by studying the sparsity of l_p minimization. Firstly, using Example 2 we show, in general, that the answer to the question above is negative. Secondly, although the answer to the question above is negative, we can prove that, for a given underdetermined linear system of equations $A_{m\times n}X = b$, there exists a constant $\gamma(A, b) > 0$ (we call it sparsity constant) such that the following conclusions hold when $p < \gamma(A, b)$.

(1) The problem (P_p) generates sparser solution as the value of p decreases (Theorem 7).

(2) Let X_p be the sparsest optimal solution to the problem (P_p). Then X_p is the unique sparsest optimal solution to the problem (P_p) under the sense of absolute value permutation (Corollary 6).

(3) Let X_1 and X_2 be the sparsest optimal solution to the problem (P_{p_1}) and problem (P_{p_2}) $(p_1 < p_2)$, respectively, and let X_1 not be the absolute value permutation of X_2. Then there exist $t_1, t_2 \in [p_1, p_2]$ such that X_1 is the sparsest optimal solution to the problem (P_t) $(\forall t \in [p_1, t_1])$ and X_2 is the sparsest optimal solution to the problem (P_t) $(\forall t \in (t_2, p_2])$ (Theorem 8).

2. The Sparsity of Underdetermined Linear System via l_p Minimization

Let $\mathcal{X}$ be the set of all solutions to the underdetermined linear systems $AX = b$. For the convenience of account, we call X_1 the absolute value permutation of X_2, which means that $(|x_{11}|, |x_{12}|, \ldots, |x_{1n}|)$ is the permutation of $(|x_{21}|, |x_{22}|, \ldots, |x_{2n}|)$, where $X_1 = (x_{11}, x_{12}, \ldots, x_{1n})^T$ and $X_2 = (x_{21}, x_{22}, \ldots, x_{2n})^T \in \mathcal{X}$.

Lemma 1 (see [17]). *The problem (P_1) may have more than one solution. Nevertheless, even if there are infinitely many possible solutions to this problem, we can claim that (1) these solutions are gathered in a set that is bounded and convex, and (2) among these solutions, there exists at least one with at most m nonzeros.*

The following example shows that, in general, it is not certain that the problem (P_p) generates sparser solution than the problem (P_1) and the problem (P_p) generates sparser solution as the value of p decreases.

Example 2. We consider the underdetermined linear system of equations $AX = b$, where

$$A = (\alpha_1, \alpha_2, \alpha_3, \alpha_4) = \begin{pmatrix} -\dfrac{20}{29} & 1 & \dfrac{31}{87} & 0 \\ 0 & 1 & \dfrac{8}{15} & 1 \\ \dfrac{60}{29} & 0 & \dfrac{463}{435} & -1 \end{pmatrix}, \tag{5}$$

$b = (1, 2, 3)^T$. By Lemma 1, the l_0-norm of the optimal solutions to the problem (P_1) are not more than 3, and hence the optimal solution is one of the following feasible solutions:

(1) $X_1 = (0, -4/27, 29/9, 58/135)^T$;

(2) $X_2 = (0.1, 0, 3, 0.4)^T$;

(3) $X_3 = (1.45, 2, 0, 0)^T$;

(4) $X_4 = (1.45, 2, 0, 0)^T$.

Furthermore, we can show that the optimal solution to the problem (P_p) $(p = 0.8, 0.95)$ is one of above feasible solutions. It is easy to calculate that

$$\|X_1\|_{0.8}^{0.8} = 3.2756,$$

$$\|X_2\|_{0.8}^{0.8} = 3.0472,$$

$$\|X_3\|_{0.8}^{0.8} = \|X_4\|_{0.8}^{0.8} = 3.0873,$$

$$\|X_1\|_{0.95}^{0.95} = 3.6502,$$
$$\|X_2\|_{0.95}^{0.95} = 3.3706,$$
$$\|X_3\|_{0.95}^{0.95} = \|X_4\|_{0.95}^{0.95} = 3.3552,$$
$$\|X_1\|_1 = 3.7999,$$
$$\|X\|_1 = 3.5,$$
$$\|X_3\|_1 = \|X_4\|_1 = 3.45. \tag{6}$$

Thus X_2 is the optimal solution when $p = 0.8$ and X_3 is the optimal solution when $p = 0.95$ and $p = 1$. However, $\|X_2\|_0 = 3$, $\|X_3\|_0 = 2$. Therefore, the problem (P_p) does not generate sparser solution than the problem (P_1) and the problem (P_p) does not generate sparser solution as the value of p decreases.

In the following, we will prove the conclusions mentioned in Introduction.

We define two functions $f(t) = \|X\|_t = (|x_1|^t + \cdots + |x_k|^t)^{1/t}$ $(t > 0)$ and $g(t) = \|X\|_t^t = |x_1|^t + \cdots + |x_k|^t$ $(t > 0)$, where $X = (x_1, \ldots, x_k)$ and $x_i \neq 0$. Then $f(t) = (g(t))^{1/t}$.

Theorem 3. *$f(t)$ is a monotone decreasing convex function and*

$$f'(t) = \frac{f(t)}{t}\left(\frac{g'(t)}{g(t)} - \ln f(t)\right). \tag{7}$$

Proof. It is easy to show that (7) holds. Without loss of generality, we assume that $|x_1| \le |x_2| \le \cdots \le |x_k|$. Because

$$\begin{aligned} f'(t) &= \frac{f(t)}{t}\left(\frac{g'(t)}{g(t)} - \ln f(t)\right) \\ &= \frac{f(t)}{t^2}\left(\frac{\sum_{i=1}^{k}|x_i|^t \ln|x_i|^t}{g(t)} - \ln g(t)\right) \\ &\le \frac{f(t)}{t^2}\left(\ln|x_k|^t - \ln g(t)\right) \le 0, \end{aligned} \tag{8}$$

$f(t)$ is monotone decreasing.

Furthermore, $f(t)$ is a convex function. In fact, we have, by the convexity of function $f(x) = x^2$,

$$\left(\frac{\sum_{i=1}^{k}|x_i|^t \ln|x_i|}{\sum_{i=1}^{k}|x_i|^t}\right)^2 \le \frac{\sum_{i=1}^{k}|x_i|^t \ln^2|x_i|}{\sum_{i=1}^{k}|x_i|^t}. \tag{9}$$

That is,

$$\left(\frac{g'(t)}{g(t)}\right)^2 \le \frac{g''(t)}{g(t)}. \tag{10}$$

Thus

$$\left(\frac{g'(t)}{g(t)}\right)' = \frac{g''(t)}{g(t)} - \left(\frac{g'(t)}{g(t)}\right)^2 \ge 0 \tag{11}$$

and hence $g'(t)/g(t)$ is monotone increasing. Since $f(t)$ is monotone decreasing, we know that $g'(t)/g(t) - \ln f(t)$ is monotone increasing. Because $f(t)/t$ is monotone decreasing, $g'(t)/g(t)$ is monotone increasing and $g'(t)/g(t) - \ln f(t) \le 0$,

$$\begin{aligned} f''(t) &= \left(\frac{f(t)}{t}\right)'\left(\frac{g'(t)}{g(t)} - \ln f(t)\right) \\ &\quad + \frac{f(t)}{t}\left(\frac{g'(t)}{g(t)} - \ln f(t)\right)' \ge 0 \end{aligned} \tag{12}$$

which implies that $f(t)$ is convex function. □

Theorem 4. *For a given underdetermined linear system of equations $A_{m\times n}X = b$, there exists a constant $\gamma > 0$ such that, for any $X_1, X_2 \in \mathcal{X}$, either $f_1'(t) = (\|X_1\|_t)' < f_2'(t) = (\|X_2\|_t)'$ or $f_2'(t) = (\|X_2\|_t)' < f_1'(t) = (\|X_1\|_t)'$ when $0 < t < \gamma$.*

Proof. Let $X^k = \{X \mid \|X\|_0 = k, X \in \mathcal{X}\}$ and $X_\beta^k = \{X \in X^k \mid$; there exists β such that $\prod_{i=1}^{k}|x_i| = \beta\}$. Clearly, we have $X^k = \cup_\beta X_\beta^k$, $\mathcal{X} = \cup_{k=1}^{n} X^k$.

Firstly, for any $X_1, X_2 \in X_\beta^k$, there exists a constant $\gamma_\beta^k > 0$ such that when $0 < t < \gamma_\beta^k$, either $f_1'(t) = (\|X_1\|_t)' < f_2'(t) = (\|X_2\|_t)'$ or $f_2'(t) = (\|X_2\|_t)' < f_1'(t) = (\|X_1\|_t)'$.

Obviously, for any given $X_1, X_2 \in X_\beta^k$, there is a positive number $\{\gamma_\beta^k\}_j$ such that when $0 < t < \{\gamma_\beta^k\}_j$, either $f_1'(t) = (\|X_1\|_t)' < f_2'(t) = (\|X_2\|_t)'$ or $f_2'(t) = (\|X_2\|_t)' < f_1'(t) = (\|X_1\|_t)'$. Hence, it suffices to show $\inf_j\{\gamma_\beta^k\}_j = \gamma_\beta^k \neq 0$. Otherwise, for an arbitrarily small positive number ε, there exists t with $0 < t < \varepsilon$, $Y_1 \in X_\beta^k$, and $Y_2 \in X_\beta^k$ such that

$$f_1'(t) = (\|Y_1\|_t)' = f_2'(t) = (\|Y_2\|_t)'. \tag{13}$$

Using (7) we obtain

$$\begin{aligned} &\frac{f_1(t)}{t}\left(\frac{g_1'(t)}{g_1(t)} - \ln f_1(t)\right) \\ &= \frac{f_2(t)}{t}\left(\frac{g_2'(t)}{g_2(t)} - \ln f_2(t)\right). \end{aligned} \tag{14}$$

That is,

$$\frac{g_1'(t)/g_1(t) - \ln f_1(t)}{g_2'(t)/g_2(t) - \ln f_2(t)} = \frac{f_2(t)}{f_1(t)}. \tag{15}$$

Since $Y_1, Y_2 \in X_\beta^k$, we have $\prod_{i=1}^{k}|y_{1i}| = \prod_{i=1}^{k}|y_{2i}| = \beta$.

Hence

$$\sum_{i=1}^{k}\ln|y_{1i}| = \sum_{i=1}^{k}\ln|y_{2i}|. \tag{16}$$

Therefore, there is a positive integer M such that

$$\sum_{i=1}^{k}\ln^M|y_{1i}| \neq \sum_{i=1}^{k}\ln^M|y_{2i}| \tag{17}$$

and, for any positive integer N with $N < M$,

$$\sum_{i=1}^{k} \ln^{N} |y_{1i}| = \sum_{i=1}^{k} \ln^{N} |y_{2i}|. \tag{18}$$

Since, for any positive integer K,

$$\begin{aligned} g_1^{(K)}(0) &= \left(|y_{11}|^t + \cdots + |y_{1k}|^t\right)^{(K)}\Big|_{t=0} = \sum_{i=1}^{k} \ln^{K} |y_{1i}|, \\ g_2^{(K)}(0) &= \left(|y_{21}|^t + \cdots + |y_{2k}|^t\right)^{(K)}\Big|_{t=0} = \sum_{i=1}^{k} \ln^{K} |y_{2i}|, \end{aligned} \tag{19}$$

we obtain, for M and N mentioned above,

$$\begin{aligned} g_1^{(M)}(0) &\neq g_2^{(M)}(0), \\ g_1^{(N)}(0) &= g_2^{(N)}(0). \end{aligned} \tag{20}$$

We assume, without loss of generality, that $g_1^{(M)}(0) < g_2^{(M)}(0)$. For the M mentioned above, (15) becomes

$$\begin{aligned} &\left[\frac{g_1'(t)/g_1(t) - \ln f_1(t)}{g_2'(t)/g_2(t) - \ln f_2(t)}\right]^{1/t^{M-1}} \\ &\quad = \left[\frac{f_2(t)}{f_1(t)}\right]^{1/t^{M-1}} = \left[\frac{g_2(t)}{g_1(t)}\right]^{1/t^{M}}. \end{aligned} \tag{21}$$

For the right of (21), we obtain

$$\begin{aligned} \lim_{t\to 0}\left[\frac{g_2(t)}{g_1(t)}\right]^{1/t^M} &= \exp\left\{\lim_{t\to 0}\frac{\ln g_2(t) - \ln g_1(t)}{t^M}\right\} = \exp\left\{\lim_{t\to 0}\frac{g_2'(t)/g_2(t) - g_1'(t)/g_1(t)}{Mt^{M-1}}\right\} \\ &= \exp\left\{\lim_{t\to 0}\frac{g_2''(t)/g_2(t) - g_1''(t)/g_1(t) - g_2'^2(t)/g_2^2(t) + g_2'^2(t)/g_1^2(t)}{Mt^{M-1}}\right\} = \cdots \\ &= \exp\left\{\frac{g_2^{(M)}(0) - g_1^{(M)}(0)}{k}\right\} > 1. \end{aligned} \tag{22}$$

And for the left of (21), we obtain

$$\begin{aligned} &\lim_{t\to 0}\left[\frac{g_1'(t)/g_1(t) - \ln f_1(t)}{g_2'/g_2(t) - \ln f_2(t)}\right]^{1/t^{M-1}} \\ &\quad = \exp\left\{\lim_{t\to 0}\frac{\ln\left(\ln g_1(t) - \left(g_1'(t)/g_1(t)\right)\times t\right) - \ln\left(\ln g_2(t) - \left(g_2'(t)/g_2(t)\right)\times t\right)}{t^{M-1}}\right\} = 1. \end{aligned} \tag{23}$$

This is a contradiction and thus when $0 < t < \gamma_\beta^k$, either $f_1'(t) = (\|X_1\|_t)' < f_2'(t) = (\|X_2\|_t)'$ or $f_2'(t) = (\|X_2\|_t)' < f_1'(t) = (\|X_1\|_t)'$.

Secondly, for any $X_1, X_2 \in X^k$, there exists a constant $\gamma^k > 0$ such that when $0 < t < \gamma^k$, either $f_1'(t) = (\|X_1\|_t)' < f_2'(t) = (\|X_2\|_t)'$ or $f_2'(t) = (\|X_2\|_t)' < f_1'(t) = (\|X_1\|_t)'$.

It suffices to show that $\inf_\beta \gamma_\beta^k = \gamma^k \neq 0$. Otherwise, for an arbitrarily small positive number ε, there is t with $0 < t < \varepsilon$, $Y_1 \in X_{\beta_1}^k$ and $Y_2 \in X_{\beta_2}^k$ $(\beta_1 \neq \beta_2)$ such that

$$f_1'(t) = \left(\|Y_1\|_t\right)' = f_2'(t) = \left(\|Y_2\|_t\right)'. \tag{24}$$

Using (7) again, we also obtain (15).

For the right of (15), we have

$$\begin{aligned} \lim_{t\to 0}\frac{f_2(t)}{f_1(t)} &= \exp\left\{\lim_{t\to 0}\frac{\ln g_2(t) - \ln g_1(t)}{t}\right\} \\ &= \frac{\prod_i |y_{1i}|}{\prod_i |y_{2i}|} \neq 1. \end{aligned} \tag{25}$$

And for the left of (15), we have

$$\begin{aligned} &\lim_{t\to 0}\frac{g_1'(t)/g_1(t) - \ln f_1(t)}{g_2'(t)/g_2(t) - \ln f_2(t)} \\ &\quad = \lim_{t\to 0}\frac{\left(g_1'(t)/g_1(t)\right)\times t - \ln g_1(t)}{\left(g_2'(t)/g_2(t)\right)\times t - \ln g_2(t)} = 1. \end{aligned} \tag{26}$$

This is a contradiction and thus $\inf_\beta \gamma_\beta^k = \gamma^k \neq 0$.

Thirdly, for any $X_1 \in X^k$, $X_2 \in X^s$, $k \neq s$, there exists a constant $\gamma^{k,s} > 0$ such that when $0 < t < \gamma^{k,s}$, either $f_1'(t) = (\|X_1\|_t)' < f_2'(t) = (\|X_2\|_t)'$ or $f_2'(t) = (\|X_2\|_t)' < f_1'(t) = (\|X_1\|_t)'$.

We assume, without loss of generality, that $\|X_1\|_0 = k < s = \|X_2\|_0$. Then

$$\lim_{t\to 0} \frac{g_1'(t)/g_1(t) - \ln f_1(t)}{g_2'(t)/g_2(t) - \ln f_2(t)}$$

$$= \lim_{t\to 0} \frac{\left(g_1'(t)/g_1(t)\right)\times t - \ln g_1(t)}{\left(g_2'(t)/g_2(t)\right)\times t - \ln g_2(t)} = \frac{\ln k}{\ln s} < 1, \tag{27}$$

$$\lim_{t\to 0} \frac{f_2(t)}{f_1(t)} = \exp\left\{\lim_{t\to 0} \frac{\ln g_2(t) - \ln g_1(t)}{t}\right\} = \infty.$$

So there is a positive number $\gamma^{k,s}$ such that when $t < \gamma^{k,s}$,

$$\frac{g_1'(t)/g_1(t) - \ln f_1(t)}{g_2'(t)/g_2(t) - \ln f_2(t)} < \frac{f_2(t)}{f_1(t)}, \tag{28}$$

which implies that

$$f_2'(t) = \left(\|X_2\|_t\right)' < f_1'(t) = \left(\|X_1\|_t\right)'. \tag{29}$$

In conclusion, we take $\gamma = \min\{\gamma^k, \gamma^{k,s} \mid k, s = 1, 2, \ldots, n\}$ and thus when $0 < t < \gamma$, for any $X_1, X_2 \in \mathcal{X}$, either $f_1'(t) = (\|X_1\|_t)' < f_2'(t) = (\|X_2\|_t)'$ or $f_2'(t) = (\|X_2\|_t)' < f_1'(t) = (\|X_1\|_t)'$. □

Obviously, for a given underdetermined linear system of equations $A_{m\times n}X = b$, there are infinitely many constants $\gamma_i > 0$ such that when $0 < t < \gamma_i$ Theorem 7 holds. The supremum of γ_i is called the sparse constant of underdetermined linear system of equations $A_{m\times n}X = b$ and denoted $\gamma(A, b)$.

Corollary 5. *Let equations $A_{m\times n}X = b$ be an underdetermined linear system. Then $f_1(t) = \|X_1\|_t$ and $f_2(t) = \|X_2\|_t$ have at most one intersection in $(0, \gamma(A, b))$, where $X_1, X_2 \in \mathcal{X}$ and X_1 is not the absolute value permutation of X_2.*

Proof. It is easy to prove that the conclusion holds by Theorems 4 and 7. □

Corollary 6. *Let X_p be the sparsest optimal solution to the problem (P_p) $(p < \gamma(A, b))$. Then X_p is the unique sparsest optimal solution to the problem (P_p) under the sense of absolute value permutation.*

Proof. Suppose that X_{p^*} is another sparsest optimal solution to the problem (P_p) and X_{p^*} is not the absolute value permutation of X_p. By Theorem 7, $\forall t \in (0, p)$, either $f_1'(t) = (\|X_p\|_t)' < f_2'(t) = (\|X_{p^*}\|_t)'$ or $f_1'(t) = (\|X_p\|_t)' > f_2'(t) = (\|X_{p^*}\|_t)'$. We suppose that $f_1'(t) = (\|X_p\|_t)' < f_2'(t) = (\|X_{p^*}\|_t)'$ and hence $\forall t \in (0, p)$ we have $f_1(t) > f_2(t)$, which implies that $\|X_p\|_0 > \|X_{p^*}\|_0$. This is a contradiction. □

Theorem 7. *The problem (P_p) generates sparser solution as the value of p decrease when $p < \gamma(A, b)$.*

Proof. If the conclusion does not hold, then there exists the optimal solutions X_1 to the problems (P_{p_1}) and the optimal solutions X_2 to the problems (P_{p_2}) satisfying $p_1 < p_2 < \gamma(A, b)$ and $\|X_1\|_0 = s > k = \|X_2\|_0$. We consider the following two cases.

(1) If $\|X_1\|_{p_1} = \|X_2\|_{p_1}$, then $\|X_1\|_{p_2} < \|X_2\|_{p_2}$ because of Corollary 5 and $s > k$. This contradicts with the fact that X_2 is the optimal solutions to (P_{p_2}).

(2) If $\|X_1\|_{p_1} < \|X_2\|_{p_1}$, then $\|X_1\|_t$ and $\|X_2\|_t$ have at least one intersection in $(0, p_1)$ because of $s > k$. Since $\|X_2\|_{p_2} \le \|X_1\|_{p_2}$, $\|X_1\|_t$, and $\|X_2\|_t$ have at least one intersection in $(p_1, p_2]$. This is contradictory to Corollary 5.

□

Theorem 8. *Let X_1 and X_2 be the sparsest optimal solution to the problem (P_{p_1}) and problem (P_{p_2}) $(p_1 < p_2 < \gamma(A, b))$, respectively, and X_1 is not the absolute value permutation of X_2. Then there exist $t_1, t_2 \in [p_1, p_2]$ such that when $p_1 \le t \le t_1$, X_1 is the sparsest optimal solution to the problem (P_t) and when $t_2 < t \le p_2$, X_2 is the sparsest optimal solution to the problem (P_t).*

Proof. Firstly, X_1 is not the optimal solution to P_{p_2} and hence $\|X_1\|_{p_2} > \|X_2\|_{p_2}$. In fact, if $\|X_1\|_{p_2} = \|X_2\|_{p_2}$, then $\|X_1\|_{p_1} < \|X_2\|_{p_1}$ by Corollary 5 and X_1 is the optimal solution to the problem (P_{p_1}). By Corollary 5 again, we have $\|X_1\|_0 < \|X_2\|_0$ which contradicts with the fact that X_2 is the sparsest optimal solutions to (P_{p_2}).

We consider the following two cases.

(1) If $\|X_1\|_{p_1} = \|X_2\|_{p_1}$, then, for any $p_2 \ge t > p_1$, X_2 is the sparsest optimal solution to the problem (P_t). Otherwise, there exists X_3 such that $\|X_3\|_t < \|X_2\|_t$ or $\|X_3\|_t = \|X_2\|_t$ and $\|X_3\|_0 < \|X_2\|_0$. If $\|X_3\|_t < \|X_2\|_t$, then $\|X_3\|_0 > \|X_2\|_0$ by Corollary 5 and $\|X_3\|_{p_1} \ge \|X_1\|_{p_1} = \|X_2\|_{p_1}$, which is contradictory to Theorem 8. If $\|X_3\|_t = \|X_2\|_t$ and $\|X_3\|_0 < \|X_2\|_0$, then $\|X_3\|_{p_1} < \|X_2\|_{p_1} = \|X_1\|_{p_1}$ by Corollary 5, which contradicts the fact that X_1 is the optimal solutions to (P_{p_1}). Therefore, we pick $t_1 = t_2 = p_1$.

(2) If $\|X_1\|_{p_1} < \|X_2\|_{p_1}$, then, by $\|X_1\|_{p_2} > \|X_2\|_{p_2}$, $\|X_1\|_t$ and $\|X_2\|_t$ have one intersection t_0 in (p_1, p_2), and hence $\|X_1\|_0 < \|X_2\|_0$. We assume, without loss of generality, that $\|X_1\|_0 + 2 = \|X_2\|_0$. Let X_3 be the sparsest optimal solution to the problem P_{t_0}. Then X_3 is not the absolute value permutation of X_2. Otherwise, we have $\|X_3\|_{t_0} = \|X_2\|_{t_0} = \|X_1\|_{t_0}$, that is, X_1 is the optimal solution to the problem P_{t_0}. Since $\|X_1\|_{p_1} < \|X_2\|_{p_1} = \|X_3\|_{p_1}$, we have $\|X_1\|_0 < \|X_3\|_0$ which contradicts the fact that X_3 is the sparsest optimal solution to the problem P_{t_0}.

If X_3 is the absolute value permutation of X_1, then $\|X_3\|_{t_0} = \|X_1\|_{t_0} = \|X_2\|_{t_0}$ and thus, by the proof of case (1), for any $p_2 \geq t > t_0$, X_2 is the sparsest optimal solution to the problem (P_t). Obviously, for any $p_1 \leq t \leq t_0$, X_1 is the sparsest optimal solution to the problem (P_t). Therefore, we pick $t_1 = t_2 = t_0$.

If X_3 is not the absolute value permutation of X_1, then $\|X_3\|_0 = \|X_1\|_0 + 1$ by Corollary 6, and there exist $t_1 \in (p_1, t_0)$, $t_2 \in (t_0, p_2)$ such that t_1 is the intersection of $\|X_3\|_t$ and $\|X_1\|_t$ and t_2 is the intersection of $\|X_3\|_t$ and $\|X_2\|_t$. By the proof above, we have that, for any $t \leq t_1$, X_1 is the sparsest optimal solution to the problem (P_t) and for any $t > t_2$, X_2 is the sparsest optimal solution to the problem (P_t). □

3. Conclusion

In this paper, the sparsity of underdetermined linear system via l_p minimization for $0 < p < 1$ has been studied. Our research reveals that for a given underdetermined linear system of equations $A_{m\times n}X = b$ there exists a sparse constant $\gamma(A, b) > 0$ such that when $p < \gamma(A, b)$, the problem (P_p) generates sparser solution as the value of p decreases and the sparsest optimal solution to the problem (P_p) is unique under the sense of absolute value permutation and if X_1 is not the absolute value permutation of X_2 where X_1 and X_2 are the sparsest optimal solution to the problems (P_{p_1}) and (P_{p_2}) $(p_1 < p_2)$, respectively, then there exist $t_1, t_2 \in [p_1, p_2]$ such that X_1 is the sparsest optimal solution to the problem (P_t) $(\forall t \in [p_1, t_1])$ and X_2 is the the sparsest optimal solution to the problem (P_t) $(\forall t \in (t_2, p_2])$.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Project 11271297, Project 11131006, Project 11201362, and EU FP7 Projects EYE2E (269118) and the National Basic Research Program of China under Project 2013CB329404.

References

[1] M.-J. Lai, "On sparse solutions of underdetermined linear systems," *Journal of Concrete and Applicable Mathematics*, vol. 8, no. 2, pp. 296–327, 2010.

[2] B. K. Natarajan, "Sparse approximate solutions to linear systems," *SIAM Journal on Computing*, vol. 24, no. 2, pp. 227–234, 1995.

[3] R. G. Baraniuk, "Compressive sensing," *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 118–124, 2007.

[4] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Review*, vol. 51, no. 1, pp. 34–81, 2009.

[5] E. J. Candès, "Compressive sampling," in *Proceedings of the International Congress of Mathematician*, vol. 3, pp. 1433–1452, European Mathematical Society, Zürich, Switzerland, 2006.

[6] R. Gribonval and M. Nielsen, "Sparse representations in unions of bases," *IEEE Transactions on Information Theory*, vol. 49, no. 12, pp. 3320–3325, 2003.

[7] R. Chartrand, "Exact reconstruction of sparse signals via nonconvex minimization," *IEEE Signal Processing Letters*, vol. 14, no. 10, pp. 707–710, 2007.

[8] R. Chartrand and V. Staneva, "Restricted isometry properties and nonconvex compressive sensing," *Inverse Problems*, vol. 24, no. 3, Article ID 035020, pp. 1–14, 2008.

[9] X. Chen, F. Xu, and Y. Ye, "Lower bound theory of nonzero entries in solutions of $l_2 - l_q$ minimization," *SIAM Journal on Scientific Computing*, vol. 32, no. 5, pp. 2832–2852, 2010.

[10] S. Foucart and M.-J. Lai, "Sparsest solutions of underdetermined linear systems via l_q-minimization for $0 < q \leq 1$," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 395–407, 2009.

[11] S. Foucart and M.-J. Lai, "Sparse recovery with pre-Gaussian random matrices," *Studia Mathematica*, vol. 200, no. 1, pp. 91–102, 2010.

[12] M.-J. Lai and J. Wang, "An unconstrained l_q Minimization with $0q \leq 1$ for sparse solution of underdetermined linear systems," *SIAM Journal on Optimization*, vol. 21, no. 1, pp. 82–101, 2011.

[13] Q. Sun, "Recovery of sparsest signals via l_q-minimization," *Applied and Computational Harmonic Analysis*, vol. 32, no. 3, pp. 329–341, 2012.

[14] Z. B. Xu, X. Chang, F. Xu, and H. Zhang, "$L_{1/2}$ regularization: a thresholding representation theory and a fast solver," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, no. 7, pp. 1013–1027, 2012.

[15] Z. B. Xu, H. L. Guo, Y. Wang, and H. Zhang, "Representative of $L_{1/2}$ regularization among $L_q(0 < q \leq 1)$ regularization: an experiments study based on phase diagram," *Acta Automatica Sinica*, vol. 38, no. 7, pp. 1225–1228, 2012.

[16] J. Zeng, S. Lin, Y. Wang, and Z. B. Xu, "$L_{1/2}$ regularization: convergence of iterative half thresholding algorithm," *IEEE Transactions on Signal Processing*, vol. 62, no. 9, pp. 2317–2329, 2014.

[17] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*, Springer, New York, NY, USA, 2010.